Human Molecular Genetics, 2014, Vol. 23, No. 18 ***4758–4769***
doi:10.1093/hmg/ddu190
Advance Access published on April 23, 2014

Mutation of senataxin alters disease-specific transcriptional networks in patients with ataxia with oculomotor apraxia type 2

Brent L. Fogel[1,*], Ellen Cho[1], Amanda Wahnich[1], Fuying Gao[2], Olivier J. Becherel[5], Xizhe Wang[1], Francesca Fike[3], Leslie Chen[1], Chiara Criscuolo[6], Giuseppe De Michele[6], Alessandro Filla[6], Abigail Collins[7,8], Angelika F. Hahn[9], Richard A. Gatti[3,4], Genevieve Konopka[10], Susan Perlman[1], Martin F. Lavin[5], Daniel H. Geschwind[1,2,4] and Giovanni Coppola[1,2]

[1]Program in Neurogenetics, Department of Neurology and [2]Department of Psychiatry, Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA, [3]Department of Pathology and Laboratory Medicine and [4]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA, [5]Radiation Biology and Oncology Laboratory, University of Queensland, UQ Centre for Clinical Research, Herston, Australia, [6]Department of Neuroscience and Reproductive and Odontostomatological Sciences, Federico II University, Napoli, Italy, [7]Department of Pediatrics and [8]Department of Neurology, Children's Hospital Colorado, University of Colorado, Denver, School of Medicine, Aurora, CO, USA, [9]Department of Clinical Neurological Sciences, Western University, London, Ontario, Canada and [10]Department of Neuroscience, University of Texas Southwestern Medical Center at Dallas, Dallas, TX, USA

Received March 15, 2014; Revised March 15, 2014; Accepted April 18, 2014

Senataxin, encoded by the *SETX* gene, contributes to multiple aspects of gene expression, including transcription and RNA processing. Mutations in *SETX* cause the recessive disorder ataxia with oculomotor apraxia type 2 (AOA2) and a dominant juvenile form of amyotrophic lateral sclerosis (ALS4). To assess the functional role of senataxin in disease, we examined differential gene expression in AOA2 patient fibroblasts, identifying a core set of genes showing altered expression by microarray and RNA-sequencing. To determine whether AOA2 and ALS4 mutations differentially affect gene expression, we overexpressed disease-specific *SETX* mutations in senataxin-haploinsufficient fibroblasts and observed changes in distinct sets of genes. This implicates mutation-specific alterations of senataxin function in disease pathogenesis and provides a novel example of allelic neurogenetic disorders with differing gene expression profiles. Weighted gene co-expression network analysis (WGCNA) demonstrated these senataxin-associated genes to be involved in both mutation-specific and shared functional gene networks. To assess this *in vivo,* we performed gene expression analysis on peripheral blood from members of 12 different AOA2 families and identified an AOA2-specific transcriptional signature. WGCNA identified two gene modules highly enriched for this transcriptional signature in the peripheral blood of all AOA2 patients studied. These modules were disease-specific and preserved in patient fibroblasts and in the cerebellum of *Setx* knockout mice demonstrating conservation across species and cell types, including neurons. These results identify novel genes and cellular pathways related to senataxin function in normal and disease states, and implicate alterations in gene expression as underlying the phenotypic differences between AOA2 and ALS4.

*To whom correspondence should be addressed at: David Geffen School of Medicine, University of California at Los Angeles, Department of Neurology, 695 Charles E Young Drive South, Gonda 1206, Los Angeles, CA 90095. Tel: +1 3108256816; Fax: +1 3102672401; Email: bfogel@ucla.edu

INTRODUCTION

Despite the continually expanding number of genes associated with cerebellar ataxia (1–4), little is known regarding the mechanisms underlying pathogenesis, which could aid the development of new therapies. System level analyses provide a broad means to identify affected cellular pathways as targets for further investigation. Here, we apply this approach to the study of ataxia with oculomotor apraxia type 2 (AOA2). AOA2 is a progressive autosomal-recessive adolescent-onset neurodegenerative disease primarily characterized by severe atrophy of the cerebellar vermis with a corresponding ataxia predominantly affecting gait, sensorimotor polyneuropathy, elevated serum alpha-fetoprotein, and, in >50% of cases, oculomotor apraxia (2,3,5,6). The disease is caused by mutation of the *SETX* (senataxin) gene on chromosome 9 which encodes a large ubiquitously expressed putative DNA/RNA helicase of 2677 amino acids (7). Studies examining the senataxin protein or the yeast ortholog Sen1p have supported a variety of cellular roles in gene expression including the regulation of RNA polymerase II binding, transcriptional termination, mRNA 3′-end maturation and pre-mRNA splicing, as well as a role in the DNA damage response through the repair of double-stranded DNA breaks (8–15). In addition to the AOA2 phenotype, dominant mutations in the *SETX* gene, distinct from those which cause ataxia, result in a juvenile form of amyotrophic lateral sclerosis designated as ALS4 (16,17).

All forms of common mutations have been observed to give rise to AOA2, including missense, nonsense, splice site and frameshift mutations as well as structural variations ranging from small insertion/deletions to larger copy number variations (5–7,18–22). In contrast, mutations leading to ALS4 are more rare and have thus far been exclusively missense variants (16,17,22), suggesting a model where loss-of-function leads to the recessive AOA2 phenotype and gain-of-function leads to the dominant ALS4 (16–18). Modified dominant phenotypes have also been suggested to be associated with in-trans rare missense variations (23,24). *SETX* is quite polymorphic and new mutations are continuously being identified. This creates a diagnostic challenge for clinicians who encounter patients with variants of unknown significance (22,25), amplified significantly by prominent phenotypic overlaps with other genetic forms of recessive cerebellar ataxia, including Friedreich ataxia, ataxia-telangiectasia and AOA1 (2,3).

Until recently, a lack of knowledge concerning the cellular functions of senataxin has prevented a comprehensive testing of its role in disease (22). Given its involvement in the regulation of transcription, we addressed this question using both *in vitro* and *in vivo* approaches. We performed genome-wide transcriptional analysis with AOA2 patient fibroblasts, identifying a set of genes whose expression is altered by mutation of *SETX*. Because of the diversity in the AOA2 and ALS4 phenotypes, we next assessed the effects of disease-specific mutations on *SETX* function by overexpressing mutant and wild-type proteins in senataxin-haploinsufficient cell lines and observed mutation-specific changes in gene expression. Translating these results *in vivo*, we confirmed that mutation-specific gene expression changes occur in the peripheral blood within a cohort of patients with AOA2. Furthermore, we demonstrate that networks of genes can be defined that are functionally associated with *SETX* in normal and disease-specific states, both in cell lines and in patient peripheral blood. Lastly, we demonstrate that these networks are not tissue- or species-restricted and are relevant to neurons, as they are found in patient fibroblasts and in the cerebellum of *Setx* knockout mice. These results support a model whereby specific pathogenic variations in senataxin function lead to altered gene expression patterns and, subsequently, to distinct clinical phenotypes.

RESULTS

Gene expression profiling identifies senataxin mutation-related changes in fibroblasts derived from an AOA2 patient

Given the known and presumed roles of *SETX* in DNA transcription and RNA processing (8–15), we hypothesized that AOA2 patients might exhibit changes in gene expression related to *SETX* mutation and sought to test this hypothesis both *in vitro* and *in vivo*. For the *in vitro* studies, we utilized cells lines derived from a patient with AOA2 and a related unaffected carrier (Table 1, Family US2). The phenotypes of AOA2 patient lymphoblast and fibroblast cell lines have both been well characterized (26,27), and we chose fibroblasts for these experiments because lymphoblast cells lines tend to show markedly reduced or absent senataxin protein, whereas fibroblast cells continue to express detectable protein (Fig. 1), useful for downstream experiments (see below).

We compared the gene expression profiles between AOA2 patient and carrier fibroblasts using genome-wide microarray analysis (Supplementary Material, File 1) and identified 1043 genes whose expression was significantly altered ($P < 0.05$). Gene ontology (GO) analysis showed altered expression in genes involved in neurogenesis, signal transduction, transcription, synaptic transmission, cell cycle regulation and sensory perception (Supplementary Material, Table S1), pathways biologically related to the observed neurological phenotype of these patients.

To confirm and extend these microarray findings, we also performed next-generation sequencing on RNA from the AOA2 patient and carrier fibroblasts (Supplementary Material, File 2) and identified differential expression of 1101 significant genes ($P < 0.05$). GO was similar to that observed for the genes identified by microarray (Supplementary Material, Table S2). To rigorously assess for genes most affected by *SETX* mutation, we performed a comparison across these two platforms, obtaining a core set of 179 genes shared between the microarray and sequencing results ($P = 2.54 \times 10^{-47}$, Supplementary Material, Fig. S1A), a significant enrichment of the two data sets (Supplementary Material, Fig. S1B). Forty-three (24%) were within the top of 20% of genes exhibiting the greatest expression differences on each platform (Supplementary Material, File 3). We examined these lists for specific genes previously reported as protein interactors of senataxin (35 genes) (9), or as disease genes involved in other forms of spinocerebellar ataxia (56 genes) or amyotrophic lateral sclerosis (104 genes) (Supplementary Material, Files 1–3), and did not observe enrichment, suggesting that pathogenic effects are not mediated by pathways directly involving these specific genes. Emphasizing that impairment of the biological function of senataxin

Table 1. Characteristics of the 12 AOA2 families included in this study

Family	Country	Mutation (mRNA)[a]	Mutation (Protein)[b]	Patients ($n = 15$)	Carriers ($n = 18$)	α-fetoprotein (ng/ml)[c]	Publication
IT1	Italy	c.3466delG homozygous	p.E1156Kfs*2	1	1	18.7–23.0	Criscuolo *et al.*[19]
IT2	Italy	**c.6109A>G** homozygous	**p.N2037D**	2	1	17.9–31.0	Unpublished
IT3	Italy	**c.6109A>G** homozygous	**p.N2037D**	1	1	17.7	Unpublished
IT4	Italy	1) c.1688A>G 2) c.6107_7199del	1) p.N563S 2) p.G2036Dfs*7	1	0	26.2	Piluso *et al.*[50]
PK1	Pakistan	c.6208+2insT homozygous	p.G2036_K2070delinsE	1	1	7.8	Fogel *et al.*[39]
CN1	Canada	1) c.5927T>G 2) **c.6654+1G>T**	1) p.L1976R 2) **p.A2183_M2218del**	1	4	87.0	Unpublished
US1	USA	1) c.5927T>G 2) c.820A>G	1) p.L1976R 2) p.M274V	1	2	61.0	Unpublished
US2	USA	1) c.5927T>G 2) c.5929C>T	1) p.L1976R 2) p.L1977F	3	2	55.6	Fogel and Perlman[18]
US3	USA	1) **c.3575_3576delAT** 2) **c.6834_6839delAACAAA**	1) **p.D1192Afs*3** 2) **p.K2278_T2279del**	1	2	76.0	Unpublished
US4	USA	1) **c.2015_2016insA** 2) **c.2724delG**	1) **p.N672Kfs*20** 2) **p.K909Nfs*7**	1	2	ND	Unpublished
US5	USA	1) c.5264delC 2) c.5833G>C	1) p.T1755Ifs*31 2) p.A1945P	1	0	29.2	Fogel and Perlman[18]
US6	USA	**c.6536T>G** heterozygous[d]	**p.I2179S**	1	2	453.0	Unpublished

[a]Based on NCBI reference sequence NM_015046.5. Novel variations are bolded.
[b]Based on NCBI reference sequence NP_055861.3. Novel variations are bolded.
[c]Normal values 0–15 ng/ml. In families with multiple affected members, values are from a single patient.
[d]A second *SETX* mutation in this family has not yet been identified.
ND, not determined.

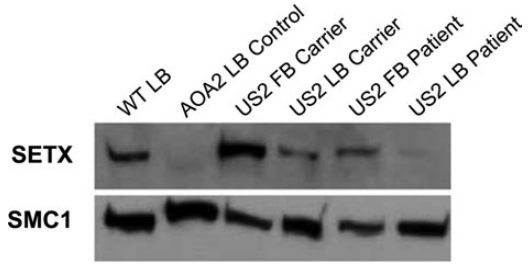

Figure 1. Expression of senataxin protein in patient and carrier fibroblasts and peripheral blood lymphoblasts. Western blotting was performed using nuclear protein from cultured lymphoblasts (LB) or fibroblasts (FB) from AOA2 patients or carriers (Family US2, Table 1) as well as a normal control (WT) using an antibody directed against senataxin (SETX). SMC1 is shown as a loading control.

causes the development of AOA2, gene ontologies of this shared list reflected involvement in neurogenesis, cell proliferation, and synaptic transmission, despite the source cell type being fibroblasts (Supplementary Material, Table S3).

Overexpression of AOA2 and ALS4 mutant senataxin differentially alters gene expression in senataxin-haploinsufficient fibroblasts

As ALS4 is extremely rare (16,17,22), patient fibroblasts could not be obtained for direct analysis. Therefore, to begin to characterize the functional effects of specific AOA2 and ALS4 mutations on the function of senataxin in the regulation of gene expression, we generated stably transfected fibroblast cell lines with wild-type and mutant forms of the senataxin protein. For studying AOA2, we chose a well-characterized founder mutation in the DNA/RNA helicase domain seen in the French-Canadian population (c.5927T>G, p.L1976R) (18,26,28) and for ALS4 we chose a previously described mutation also found in the DNA/RNA helicase domain (c.6407G>A, p.R2136H) (16,22). Because both loss-of-function and gain-of-function mechanisms have been postulated as contributing to senataxin-related disease (7,16,18),we chose an AOA2 carrier fibroblast cell line to generate these stable transfections. Because the carrier line is haploinsufficient, and thus expresses a reduced baseline level of normal functional senataxin protein (Fig. 1), both increased and decreased changes in senataxin-related gene expression could be observed upon overexpression of the mutant proteins. In addition to the mutant constructs, as a control, we also tested the effects of overexpressing a previously identified alternatively spliced form of *SETX* (hereafter referred to as 23B) which modifies the protein's DNA/RNA helicase domain (21) and represents 8% of the *SETX* transcripts expressed in the cerebellar vermis (Supplementary Material, Fig. S2).

To minimize confounding changes in gene expression resulting from our experimental procedures, all data were first normalized to baseline expression in the untransfected carrier (control) line. Gene expression in cells stably transfected with wild-type senataxin protein was compared with cells transfected with either AOA2 or ALS4 mutant proteins, or the protein from the *SETX* 23B alternatively spliced isoform. We identified genes with significant expression changes versus the control line and

determined whether each showed a similar change, an opposite change, or no change relative to the wild-type transfected senataxin line (Supplementary Material, File 4). If the expression of the gene in question changed significantly in just the ALS4 or AOA2 lines or in opposition with wild-type transfected senataxin, we considered that gene to exhibit a disease-specific change. In total, we identified 113 genes with ALS4-specific expression changes ($P < 0.05$) and 127 genes with AOA2-specific expression changes ($P < 0.05$). In contrast, only 30 genes showed expression changes specific to the control *SETX* 23B isoform ($P < 0.05$) (Supplementary Material, File 4). The ALS4-specific genes showed only a slight enrichment for differentially expressed genes from the AOA2 patient fibroblasts (11 genes, $P = 2.49 \times 10^{-2}$) and only a marginally higher enrichment was seen with the AOA2-specific genes (15 genes, $P = 1.60 \times 10^{-3}$), likely due the influence of endogenous wild-type senataxin protein within these cells. As before, we did not observe significant overlap between these ALS4- or AOA2-specific genes and sets of specific genes involved in spinocerebellar ataxia (56 genes), amyotrophic lateral sclerosis (104 genes), or as protein interactors of senataxin (35 genes) (9) (Supplementary Material, File 5). GOs were similar to those in the original patient fibroblast analysis, reflecting involvement in cell cycle regulation (AOA2-specific genes), cellular metabolism (ALS4-specific genes) and synaptogenesis (23B-specific genes) (Supplementary Material, File 5). Although the AOA2- and ALS4-specific sets did show a highly significant overlap of 18 genes (8%, $P = 7.03 \times 10^{-21}$, with 4 genes (1%) also shared with the 23B set) (Supplementary Material, File 5), likely reflecting constitutive *SETX* functions, 92% of the genes were distinct to one set or the other. Overall, these data demonstrate that disease-specific mutations in the *SETX* gene cause differential transcriptional changes within cells.

Weighted gene co-expression network analysis reveals that genes showing altered expression with AOA2 or ALS4 *SETX* mutations are part of both disease-specific and shared functional networks

To better understand the systems level organization of the transcriptional changes occurring in these fibroblasts as a result of mutation of *SETX*, we performed weighted gene co-expression network analysis (WGCNA) to identify modules of biologically related genes (29–31). To identify the modules most related to senataxin function, we assessed the module eigengene relationship with *SETX* mutation status, identifying four modules that showed the largest difference between the *SETX* mutations and wild-type, designated as the green, lightcyan, skyblue and yellow modules, respectively (Fig. 2, Supplementary Material, File 6). We examined whether there was a statistically significant enrichment of genes identified as ALS4- or AOA2-specific in these modules and found ALS4-specific genes to be enriched in the skyblue (13 genes, $P = 3.10 \times 10^{-17}$), green (13 genes, $P = 8.53 \times 10^{-9}$) and yellow (13 genes, $P = 6.44 \times 10^{-6}$) modules and AOA2-specific genes to be enriched in the green (15 genes, $P = 4.12 \times 10^{-10}$) and lightcyan (7 genes, $P = 8.01 \times 10^{-6}$) modules (Supplementary Material, File 6). The green and skyblue modules showed the most significant enrichment for AOA2- and ALS4-specific genes, respectively (Fig. 2, Supplementary Material, File 6). As in the previous experiments, none of the modules were enriched for specific genes involved in spinocerebellar ataxia (56 genes), amyotrophic lateral sclerosis (104 genes) or as protein interactors of senataxin (35 genes) (9) (Supplementary Material, File 6). GO analysis of the less enriched modules identified genes associated with organ development (lightcyan module) and cellular metabolism and transcription (yellow module) (Supplementary Material, File 6). The skyblue module, which contains 67 members, was associated with protein and RNA processing and enriched for ALS4-specific genes, including two of the four most connected (hub) genes, *DDX21* and *HNRPK*, both known regulators of RNA metabolism (Fig. 2A; Supplementary Material, File 6). The green module, which contains 300 members, was enriched for both ALS4- and AOA2-specific genes, highlighting a functional overlap between the two disease gene sets (Fig. 2B; Supplementary Material, File 6). GO analysis supported associations with cell cycle regulation and with DNA maintenance including repair, replication and transcription (Supplementary Material, File 6). Together, these results further support and extend several functions previously suggested or attributed to senataxin involving aspects of RNA processing, DNA maintenance and transcription (7,9,11,12,14–16,27). In summary, while the ALS4- and AOA2-specific gene sets identify disease-specific functional gene networks, they also converge on similar biological pathways more generally related to senataxin cellular function.

Gene expression profiling in patients identifies a transcriptional signature for AOA2 in peripheral blood

Given the existing diagnostic challenges associated with verifying novel pathogenic *SETX* variants, we asked whether relevant changes in gene expression could be observed in the peripheral blood of patients with AOA2. To test this, we utilized a cohort of 15 patients and 18 carriers from 12 distinct families representing 4 different countries of origin (Table 1), the largest group of AOA2 patients studied in this manner to date. Whole transcriptome microarray analysis was performed on patient and carrier RNA obtained from the peripheral blood. To reduce the risk of family-specific gene expression changes influencing the results, data within each family were averaged and patients were normalized to carriers. To maximize the identification of AOA2-specific changes, five families were selected for initial comparison based on most diverse countries of origin (Table 1) representing the USA (US1, US2; 2 families), Canada (CN1; 1 family), Italy (IT1; 1 family) and Pakistan (PK1; 1 family). A list of genes with the most significantly altered expression across all five families was identified, resulting in a set of 204 genes ($P < 0.01$) (Supplementary Material, File 7). Similar to what was seen in the patient fibroblast cell lines, GO showed increased expression of genes involved in RNA processing and decreased expression of genes involved in cellular metabolism (Supplementary Material, File 7). No enrichment was seen for specific genes implicated in spinocerebellar ataxia (56 genes), amyotrophic lateral sclerosis (104 genes), or as protein interactors of senataxin (35 genes) (9). Additionally, no enrichment was seen with the AOA2 genes from patient fibroblasts or with the ALS4- or AOA2-specific gene sets identified in the cell lines derived from them (Supplementary Material, File 7). Variability was also observed in the relative levels of gene expression across individuals when using

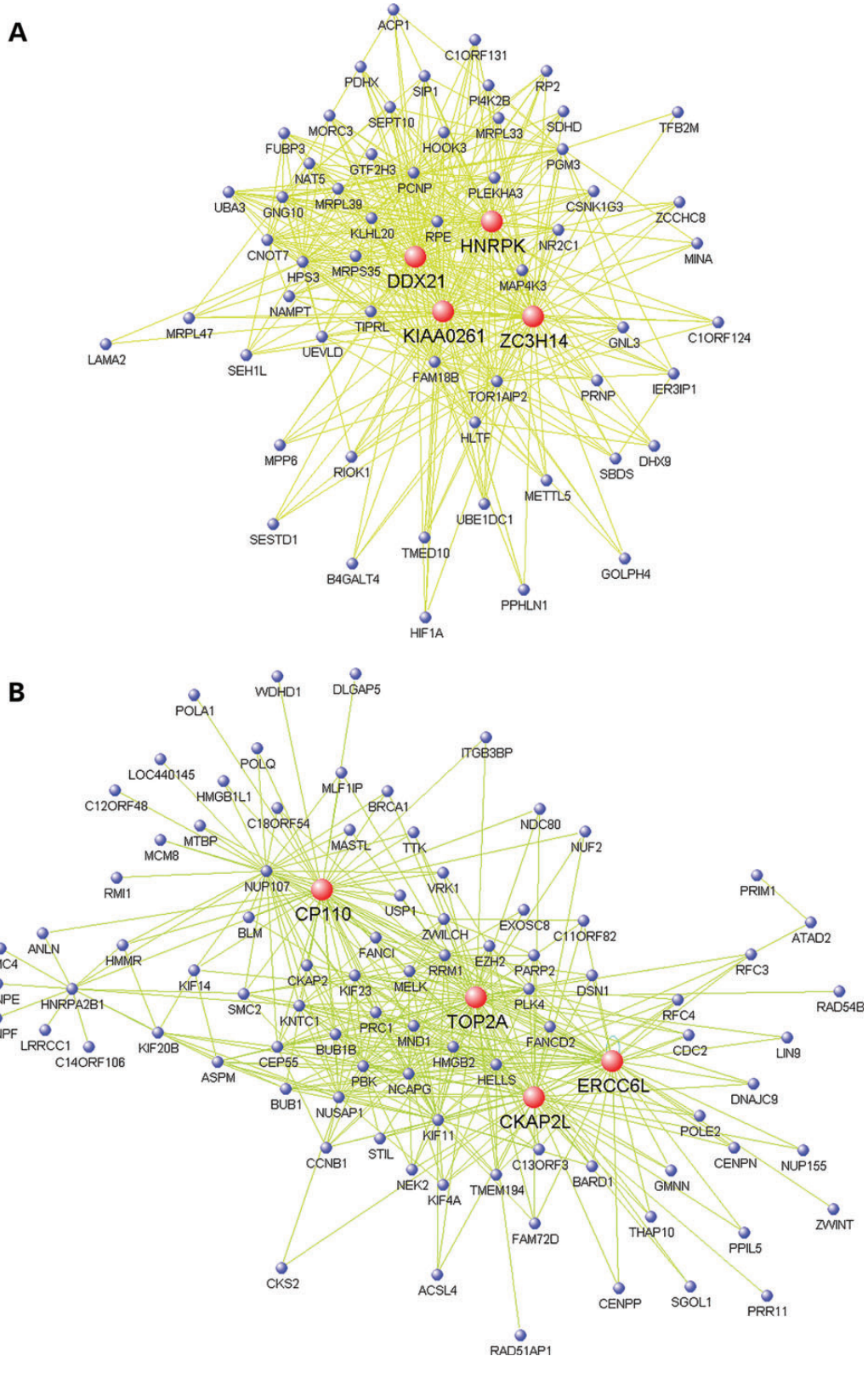

Figure 2. Weighted gene co-expression network analysis (WGCNA) in senataxin-haploinsufficient fibroblasts overexpressing mutant forms of senataxin identifies molecular pathways related to biological function. WGCNA was performed to identify groups of genes (modules) whose collective changes in gene expression most correlate with senataxin mutation, reflecting biologically relevant relationships. For clarity, only the most highly connected module members are shown. Genes with the highest connectivity (i.e. hubs) are indicated in red. (**A**) Skyblue module. (**B**). Green module.

a classifier algorithm to assess for the presence or absence of the AOA2 transcriptional signature in the 15 AOA2 patients (Table 1) and a control set of 52 individuals, resulting in a modest sensitivity of 0.67 and a specificity of 0.85 (Supplementary Material, File 8), suggesting a systems-based approach might provide a more comprehensive view of the expression changes occurring across patients.

Weighted gene co-expression network analysis identifies modules specific to AOA2 and senataxin function in patient peripheral blood

As our data indicated that the relative expression of AOA2-specific genes varies greatly across individuals, we looked to see whether we could identify patterns of correlated expression within these genes that could define AOA2 status. Using WGCNA, we compared affected patients to carriers across all our AOA2 families (Table 1) and identified four modules which correlated with disease status, termed black, blue, turquoise and yellow, respectively (Supplementary Material, File 9). GO indicated these modules were associated with RNA processing (black, yellow and turquoise modules), apoptosis (black, yellow and blue modules), translation and protein metabolism (yellow and turquoise modules), cellular metabolism (blue module), immune function (black and turquoise modules), mitochondrial function (yellow module) and DNA repair, replication and transcription (turquoise module) (Supplementary Material, File 9), consistent with known and proposed *SETX* functions (7,9,11,12,14–16,27). The blue module (1837 genes) was found to be highly enriched for AOA2-specific genes from patient peripheral blood (114/204 genes, $P = 1.44 \times 10^{-66}$; Supplementary Material, File 9), including three of the five hub genes (*CDC34, DPM2, SEMA6B*; Fig. 3A), highlighting this module as best defining the AOA2 transcriptional signature. The turquoise module (1937 genes) was also enriched for AOA2-specific genes from the peripheral blood (57/204 genes, $P = 1.11 \times 10^{-14}$; Supplementary Material, File 9) as well as for ALS4-specific genes from the fibroblast cell lines (30/113 genes, $P = 7.53 \times 10^{-8}$; Supplementary Material, File 9), and mildly enriched for senataxin protein interactors (9) (8/35 genes, $P = 1.27 \times 10^{-2}$; Supplementary Material, File 9) suggesting this to be a module more generally related to senataxin function (Fig. 3B). These results demonstrate that both mutation-specific and generalized senataxin functional gene expression networks can be detected *in vivo* in the peripheral blood of AOA2 patients.

AOA2 and senataxin functional WGCNA modules from patient peripheral blood are disease-specific, preserved across cell type and species, and are relevant to neurons

Given that our prior studies all involved non-neuronal cell types, we questioned whether these observed gene expression changes were cell-type specific or universally applicable to neurons, as predicted based on the ubiquitous expression of senataxin. To test this at a systems level, we first compared the WGCNA networks from patient peripheral blood (Supplementary Material, File 9) to the WGCNA networks derived from the transfected fibroblast cell lines (Supplementary Material, File 6). We observed that the key blood modules, black, blue, turquoise and yellow, were all significantly preserved, with the turquoise module being the most preserved overall between the networks (Fig. 3C). Significant overlap was also noted between the members of the turquoise blood module and the green (80/300 genes, $P = 1.04 \times 10^{-18}$), skyblue (28/67 genes, $P = 1.42 \times 10^{-12}$) and yellow (214/537 genes, $P = 6.97 \times 10^{-84}$) modules from the fibroblast cell lines (Supplementary Material, File 10), demonstrating these affected genetic pathways to be universal across these two cell types.

To validate these observations in neurons, we examined gene expression in the previously reported *Setx* knockout (KO) mice (15). In the cerebellum, we identified 2091 genes ($P < 0.05$) differentially expressed between KO and WT animals (Supplementary Material, File 11). Consistent with our findings in human cells, we saw no enrichment for specific genes implicated in spinocerebellar ataxia (56 genes), amyotrophic lateral sclerosis (104 genes) or as protein interactors of senataxin (35 genes) (9) and, although GO analysis was similar, there was no significant enrichment for specific genes from patient fibroblasts, ALS4- or AOA2-specific gene sets identified in the transfected cell lines, or genes from AOA2 patient peripheral blood (Supplementary Material, File 11). We next utilized a systems approach, comparing the WGCNA networks from patient peripheral blood (Supplementary Material, File 9) to WGCNA networks derived from *Setx* KO mouse cerebellum (Supplementary Material, File 12), and found significant preservation of the turquoise and blue modules, the two modules most enriched for the AOA2 transcriptional signature from patient peripheral blood (Fig. 3C). This effect was only observed in the cerebellum and not whole brain (data not shown). There was also a significant overlap between the members of the turquoise blood module and the knockout-related greenblue module from the *Setx* KO cerebellum (88/589 genes, $P = 3.96 \times 10^{-6}$) (Supplementary Material, File 10). Members of the greenblue cerebellar module also moderately overlapped with the green (14/300 genes, $P = 4.44 \times 10^{-2}$) and yellow (23/537 genes, $P = 3.02 \times 10^{-2}$) modules from the fibroblast cell lines (Supplementary Material, File 10). The greenblue module contains 589 members and GO analysis reflects the above relationships, indicating functions in gene transcription, RNA processing, brain development, DNA repair and cellular metabolism (Supplementary Material, Fig. S3 and File 12). These results confirm, at a systems level, that the genetic networks directly related to senataxin function and disease are conserved across species as well as within patient cells and, most significantly, are relevant to neurons.

To assess the specificity of this observation, we compared the WGCNA networks from AOA2 patient peripheral blood to the networks derived from the peripheral blood of a cohort of patients with Friedreich ataxia (8 carriers and 24 patients) (Supplementary Material, File 8), the most common phenotypically overlapping genetic cerebellar ataxia. Since both networks were derived from peripheral blood, preservation of all modules was seen as expected (Supplementary Material, File 8). However, correlation of the turquoise and blue models with disease status, as seen in the AOA2 cohort, was not found in the Friedreich ataxia samples ($P = 0.59$ and $P = 0.40$, respectively) (Supplementary Material, File 8). In contrast, the preserved turquoise model in the cerebellum of the *Setx* knockout mice directly correlated with knockout status ($P = 0.036$) (Supplementary Material, File 8). This confirms that these conserved

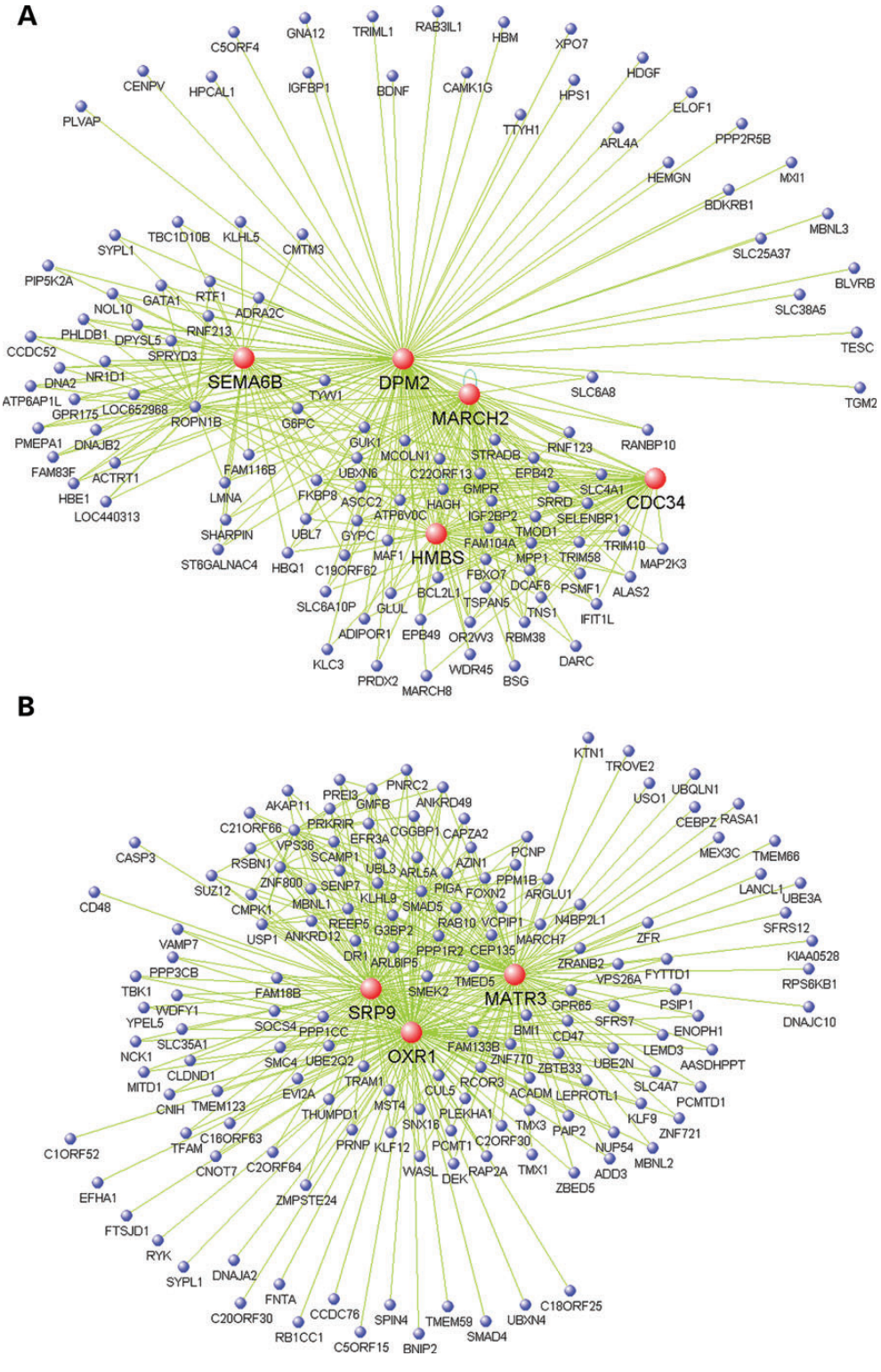

Figure 3. WGCNA in peripheral blood identifies a transcriptional signature for senataxin dysfunction in AOA2 patients. WGCNA was performed to identify gene modules whose collective changes in gene expression most correlate with AOA2 disease status. For clarity, only the most highly connected module members are shown. Genes with the highest connectivity (i.e. hubs) are indicated in red. (**A**) Blue module. (**B**). Turquoise module. (**C**) Modules from the peripheral blood weighted gene co-expression network (Supplementary Material, File 9) are shown in comparison with the network derived from human transfected fibroblast cell lines (Supplementary Material, File 6) or *Setx* knockout mouse cerebellum (Supplementary Material, File 12). Results are graphed by number of module members versus preservation score (Zsummary) calculated using the module preservation function of the WGCNA software package. Zsummary scores of 10 or greater (dashed line) indicate strong preservation, scores between 10 and 2 show moderate preservation, and scores less than 2 (dotted line) show weak or no preservation.

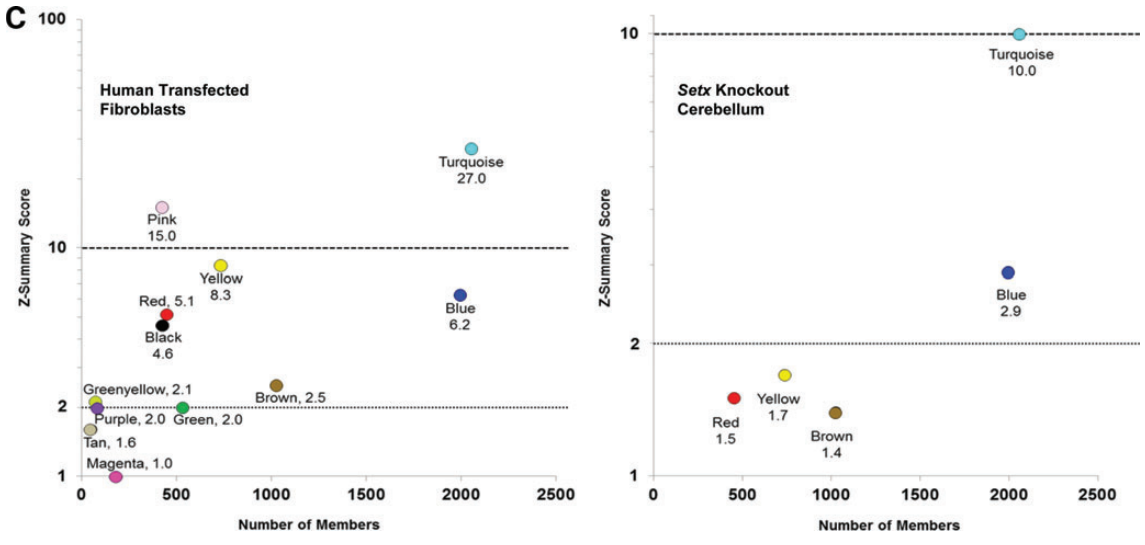

Fig. 3. *Continued*

senataxin genetic networks are specific to AOA2 and the loss of senataxin function.

DISCUSSION

Here, we sought to identify changes in gene expression associated with mutation of the senataxin protein, implicated in a diverse array of functions related to gene expression including regulating RNA polymerase II binding, transcriptional termination, mRNA 3′-end maturation, pre-mRNA splicing and the repair of double-stranded DNA breaks (8–15). Since senataxin is ubiquitously expressed, we were able to address this question both *in vitro* and *in vivo*. To study gene expression *in vitro*, we derived fibroblast cells lines from a patient with AOA2 and identified a core set of 179 differentially expressed genes using two different, yet complimentary, methods, microarray and RNA sequencing. Despite using data derived from fibroblasts, it is notable that the identified genes are primarily associated with neurogenesis and related functions (Supplementary Material, Tables S1–3), consistent with the neurological phenotypes associated with mutation of this protein. Examination of these genes revealed further interesting relationships which may explain certain clinical characteristics. As one example, it was particularly interesting that a gene showing the greatest increase in expression using both methods was *PSG4* (Supplementary Material, File 3), a member of the pregnancy-specific glycoprotein and carcinoembryonic antigen gene families (32). Given the elevated blood levels of alpha-fetoprotein, also a tumor marker, already seen in AOA2 patients, this suggests that the presence of such antigens is likely due to alterations in senataxin-related gene expression, consistent with the lack of any observed elevations in cancer risk within this patient population (2,5,6).

Of further clinical note, in the AOA2 patient fibroblasts, there were three ataxia and five ALS genes detected by microarray (Supplementary Material, File 1) and four ataxia and eight ALS genes detected by RNA sequencing (Supplementary Material, File 2), with two of the ataxia genes, *PPP2R2B* and *SLC1A3,* also being present in the shared core gene list (Supplementary Material, File 3). In addition, microarray analysis detected changes in expression of *ATM*, a DNA repair gene whose mutation causes ataxia-telangiectasia, a disease with a clinical phenotype strikingly similar to AOA2 (2). Whether altered expression of these disease genes is significant to the AOA2 phenotype will require further investigation. It is also interesting to note that *TOP2A,* one of the hubs of the green module identified in fibroblasts overexpressing mutant senataxin (Fig. 2B), has recently been shown to be a target for the ATM protein following DNA damage (33).

We observed that overexpression of mutant forms of senataxin in senataxin-haploinsufficient fibroblasts results in altered gene expression and that a portion of these changes are specific to the class of mutation introduced (AOA2 versus ALS4). This suggests that different senataxin mutations disrupt function in different ways leading to alternate patterns of gene expression (Fig. 4), a novel finding not previously reported for allelic disorders. Because many of the mutations that cause AOA2 are unstable and/or truncating (6,18,21), it is predicted that these mutations abolish function, consistent with a recessive disorder, whereas ALS4 missense mutations likely modify those functions and hence are dominant (16,18,21). This is supported by observations that an intact DNA/RNA helicase domain is required to produce the ALS4 phenotype (21), and is further supported by our data which show enrichment of ALS4-specific genes in WGCNA modules related to general senataxin function in both fibroblasts and peripheral blood (green module, Fig. 2B, Supplementary Material, File 6; turquoise module, Fig. 3B, Supplementary Material, File 9), as well as module preservation between the two cell types (Fig. 3C). It is likely that some of these effects are mediated through the regulation of transcription and other regulatory factors, as indicated by an enrichment for several known binding sites within the core genes identified in AOA2 patient fibroblasts (data not shown). Interestingly, the most significantly enriched factor is myc-associated zinc-finger protein (*MAZ*) (data not shown), which has been found to be a target of paraneoplastic antibodies causing cerebellar degeneration (34). Further supporting this, it has been recently shown that AOA2- but not ALS4-specific mutations in the N-terminus of senataxin affect

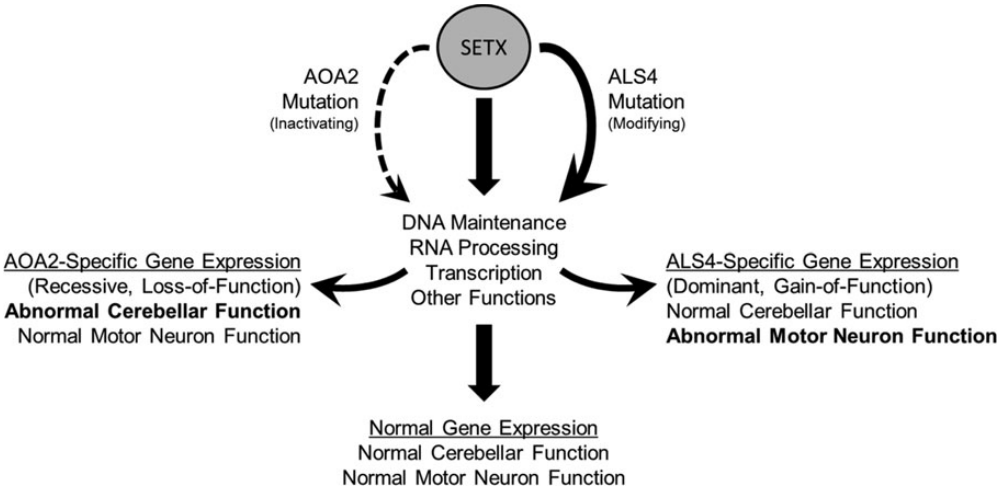

Figure 4. Model of senataxin (SETX) function in normal and disease states.

sumoylation of the protein and subsequent interaction with the exosome via association with Rrp45 (EXOSC9) (35). Notably, EXOSC9 was found in our microarray analysis of patient fibroblasts (Supplementary Material, File 1). Determining whether the C-terminal AOA2 and ALS4 mutations studied here affect senataxin function directly or indirectly deserves future study.

The identification of a molecular gene expression signature *in vivo* in the peripheral blood of AOA2 patients is consistent with findings from other neurological diseases related to mutation of broadly expressed proteins (36,37). The variability we detect among individual AOA2 patients (Supplementary Material, File 8) is likely due to either differences across genetic backgrounds or differences in effect between mutations. Supporting this, three patients and two carriers of Italian descent with the identical *SETX* genotype (families IT2 and IT3; homozygous p.N2037D) were consistently misclassified by our algorithm (Supplementary Material, File 8) yet were distinguishable as AOA2 patients or carriers by examining gene co-expression patterns (Supplementary Material, File 9). Because of this, the blue and turquoises modules (Fig. 3, Supplementary Material, File 9), both enriched for AOA2 transcriptional signature genes and preserved in the cerebellum of *Setx* knockout mice (Fig. 3C), may be useful for diagnostic assessment of variants of unknown significance in suspected clinical cases of AOA2. This is supported by the observation that these modules do not correlate with disease in the blood of Friedreich ataxia patients but, in the case of the turquoise module, do correlate with *Setx* knockout in mice (Supplementary Material, File 8). This may become particularly relevant given the highly polymorphic nature of *SETX* and the number of novel variants continually being identified, especially as clinical exome sequencing for neurological disorders becomes more prevalent (6,21,25,38,39).

Lastly, the identification of gene networks functionally related to senataxin expands the current knowledge of its cellular role and presents avenues for future investigations into its detailed function in normal cells and in disease. The turquoise model is of particular interest in this regard, as it is enriched for AOA2 genes from peripheral blood and ALS4-specific genes from patient fibroblasts (Supplementary Material, File 9) and is the most preserved module between peripheral blood and both the transfected human fibroblast and *Setx* knockout mouse cerebellum networks (Fig. 3C), where it also correlates with knockout status (Supplementary Material, File 8). Furthermore, it contains a number of factors already known to be associated with senataxin functionally, including eight interacting proteins (9) (Supplementary Material, File 9). The module also contains *BRCA1*, recently shown to be involved in senataxin-mediated meiotic sex chromosome inactivation in mice (15), as well as the genes *DDX5, CRY1, PER2* and *XRN2*, whose products are involved with senataxin in transcriptional termination (11–13). Other pathways and factors within this model may pose excellent targets for future studies of senataxin biological function and/or be potential therapeutic targets in disease. Two interesting candidates are the hub genes *OXR1,* previously shown to be involved in cerebellar-dependent motor learning (40) and oxidative stress-induced cerebellar neurodegeneration (41), and *MATR3,* a target for the ATM protein involved in the regulation of gene expression via mRNA stabilization (42). Further investigation of these networks will likely identify additional specific cellular pathways whose disruption through mutation of *SETX* contribute to the distinct clinical phenotypes seen in AOA2 and ALS4.

MATERIALS AND METHODS

SETX gene nomenclature and isoform analysis

Exon numbering in this report is based on the canonical transcript of the *SETX* gene (NM_015046.5). Exon 1 is designated as the exon containing the ATG start codon, the first base of which is position 1. Using this nomenclature, the *SETX* coding region spans 24 exons with exons 12–24 (bases 1931 to 2456) encoding a Superfamily I DNA/RNA helicase domain.

Plasmid construction, cell culture and stable transfection

Full-length cDNA representing the complete *SETX* coding sequence was derived from a larger cDNA construct (Open BioSystems, Thermo Fisher Scientific) and cloned into pENTR-D-TOPO (Invitrogen, Life Technologies) by PCR.

Point mutations were introduced using the QuikChange Lightning Multi Site-Directed Mutagenesis kit (Stratagene, Agilent Technologies) per the manufacturer's instructions. The original *SETX* cDNA contained four coding sequence variants that differed from the canonical *SETX* transcript (NM_015046.5) and these were mutagenized to match wild-type sequence creating the *SETX*wt isoform which was then used to generate the AOA2 and ALS4 variants. The 87 bp Exon 23b was added by multiplex PCR to a *PstI* fragment of the *SETX* cDNA and cloned back into the full-length wild-type cDNA to create the *SETX*+23b isoform. For expression, cDNAs were subcloned using Gateway cloning (Invitrogen, Life Technologies) into pDEST26 (Invitrogen, Life Technologies) which adds an N-terminal His-tag. All *SETX* inserts were fully sequenced after subcloning into pDEST26 to confirm all wild-type and mutant sequences.

For transfection, pDEST-*SETX* plasmids were linearized with *AhdI* (New England BioLabs) and introduced using an Amaxa Nucleofector (Lonza) per the manufacturer's instructions. Positive selection was performed using 200 μg/ml G418 (Invitrogen, Life Technologies). Two biological replicates were produced for each stable cell line and handled independently.

Molecular characterization of AOA2 families

Patients were selected based on the presence of the specific rare clinical phenotype associated with AOA2 (including data from neurological examinations, neuroimaging, electrophysiological studies and biomarker testing) (2,3,5,6) and the finding of known or suspected disease-causing sequence variations in the *SETX* gene that segregate with disease. The probands for 10 of the 11 families tested showed elevated levels of alpha-fetoprotein (Table 1). Seven previously unreported sequence variations were identified among six families, including a novel splicing variant (Supplementary Material, Fig. S4). One variant, p.N2037D, was present in two unrelated families of Italian origin (Table 1). These novel variants were collectively predicted to be deleterious because they either caused premature termination of the protein (families US3 and US4), deleted residues within the DNA/RNA helicase domain (families CN1 and US3), or resulted in missense variants not present in either the 1000 genomes database (http://www.1000genomes.org/) or the National Institutes of Health Lung and Blood Institute (NHLBI) Exome Variant Server (http://evs.gs.washington.edu/EVS/) and, additionally, were bioinformatically predicted as damaging using the PolyPhen-2 software (http://genetics.bwh.harvard.edu/pph2/index.shtml) (43) (families IT2, IT3 and US6). Carrier individuals were available for 10 of the 12 families. In one family (US6), only a single heterozygous missense sequence variant was identified and a screen for copy-number variation was negative (data not shown). This patient was also screened for other potential pathogenic variation using whole exome sequencing, but this was negative as well (data not shown). Based on consistent clinical phenotype (Table 1) and a lack of any other identifiable molecular etiology, this patient was included in the cohort and is suspected of harboring an as-yet-unidentified noncoding mutation on the second SETX allele. All novel *SETX* variants reported here are deposited in the UCLA Neurogenetics *SETX* Database (http://149.142.212.78/LOVD/home.php?select_db=SETX).

Patient cell lines, mouse tissues and preparation of RNA and protein

Skin biopsies were obtained from a patient with AOA2 (Family US2, Table 1) and a related carrier (heterozygous c.5927T>G, p.L1976R) (18) and fibroblasts were grown in DMEM with 10% fetal bovine serum and 1% penicillin streptomycin (Life Technologies). Cells were harvested and RNA was extracted using the miRNeasy Mini Kit (Qiagen) per the manufacturer's recommended instructions. The NE-PER kit (Pierce, Thermo Fisher Scientific) was used to isolate nuclear protein per the manufacturer's instructions. Protein concentrations were determined using either Bradford (Bio-Rad) or BCA (Pierce, Thermo Fisher Scientific) assays.

RNA from normal human cerebellar vermis (44) and quantitative real-time PCR (45) were as previously described. Lymphoblast cell lines were prepared using standard methods (26) and nuclear protein prepared as above. Murine tissues and RNA were prepared as previously described (15).

Preparation and analysis of DNA and RNA from peripheral blood

Consent was obtained to extract DNA for genetic analysis and all patients were provided genetic counseling. Peripheral blood cells were obtained and genomic DNA was prepared using Autopure LS (Qiagen) according to the manufacturer's instructions. Sequencing of *SETX* coding exons and the neighboring intronic junctions was performed using a 3730 DNA Analyzer (Applied Biosystems, Life Technologies). Primers utilized for sequencing were as described in Criscuolo *et al*. (19). All study methods were approved by the Institutional Review Board of the University of California, Los Angeles (UCLA).

Total RNA was extracted from the peripheral blood using the PAXgene Blood RNA Kit (Qiagen) per the manufacturer's instructions. RNA quality control was performed using an Agilent Bioanalyzer 2100 (Agilent Technologies) and reverse transcription-PCR was performed as described previously (21). All PCR products were confirmed by sequencing as above. All primer sequences are available on request. Gel images were captured using a BioSpectrum AC Imaging System (UVP).

Immunoblotting and antibodies used

Immunoblotting was performed using standard protocols (45). Primary antibodies were against SETX (mouse polyclonal, 1:1000) and SMC1 (rabbit polyclonal, 1:1000, Novus Biologicals). The secondary antibody used was goat anti-mouse horseradish peroxidase (1:4000–1:10 000, Millipore).

RNA sequencing and data analysis

Five micrograms of total RNA from each of two biological replicates of each fibroblast cell line was treated with the RiboMinus Eukaryote kit (Invitrogen, Life Technologies) to remove ribosomal RNA. Eighty nanograms of ribo-depleted RNA was used in fragmentation and 50 ng of fragmented RNA was used to generate barcoded libraries using the SOLiD Whole Transcriptome Analysis Kit and the SOLiD Transcriptome Multiplexing Kit according to the manufacturer's protocol (Applied Biosystems,

Life Technologies). The four barcoded samples were run together and sequenced at a 50 bp single-end read scale using a SOLiD 3.5 system (Applied Biosystems). Unique sequence reads were aligned to the human genome (hg19, NCBI 37) using Bowtie (46) with a 25 bp seed and three mismatch allowance. Biological replicate counts were pooled for gene expression analysis. Statistical comparisons of differential gene expression were made using a negative binomial distribution model with the DESeq software package (47). A cut-off threshold of $P < 0.05$ was used for significance. Validation was performed on a random sample of genes using quantitative reverse transcription-PCR, as previously described (Supplementary Material, Fig. S5) (45). GOs were determined using the DAVID Bioinformatic resources (NIAID, NIH). Statistical comparisons to other gene lists were made using hypergeometric probability.

Microarray analysis

Total RNA samples were analyzed using either Illumina HumanRef-8 v3.0, HumanHT-12 v4.0 or MouseRef-8 v2.0 expression arrays per the manufacturer's instructions. The human arrays share 24 525 common probes for comparison. Data were filtered for batch effects and biological replicate counts were pooled for gene expression analysis. Statistical comparisons for differential gene expression were made using the limma software package (48) and threshold for significance was set at $P < 0.05$. GOs were determined using the DAVID Bioinformatic resources (NIAID, NIH). Statistical comparisons to other gene lists were made using hypergeometric probability. Validation was performed on a random sample of genes using quantitative reverse transcription-PCR (Supplementary Material, Fig. S5) (45). WGCNA was performed as previously described (45).

All microarray and RNA sequencing data are deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/).

NOTE ADDED IN PROOF

Mutation of *MATR3*, a hub gene found in the turquoise module in the peripheral blood of patients with AOA2 (Fig. 3B), was recently reported as a rare genetic cause of familial amyotrophic lateral sclerosis (49).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

The authors thank Dr Nagesh Rao for assistance with fibroblast preparation, Dr Maduri Hegde for assistance with *SETX* copy number variation analysis, and Jessica Lane and Ji Yong Lee for technical assistance.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the National Institute of Mental Health (K08MH86297 to B.L.F. and R00MH090238 to G.K.), the National Institute for Neurological Disorders and Stroke (R01NS082094 to B.L.F.) and the National Ataxia Foundation (Young Investigator Award to B.L.F.). We acknowledge the support of the NINDS Informatics Center for Neurogenetics and Neurogenomics (P30 NS062691), the UCLA Intellectual and Developmental Disabilities Research Center (IDDRC, PC HD004612) and the Muscular Dystrophy Association (grant #186994 to G.C.). We also thank the Italian Ministry of Education, University, and Research (PRIN 2010-2011 20108WT59Y-007 to G.D.M.). Dr Fogel gratefully acknowledges the generous support through donations to the University of California by the De Mint Family.

REFERENCES

1. Fogel, B.L. and Perlman, S. (2011) Gálvez-Jiménez, N. and Tuite, P.J. (eds), In *Uncommon Causes of Movement Disorders*. Cambridge University Press, Cambridge, NY, pp. 198–216.
2. Fogel, B.L. and Perlman, S. (2007) Clinical features and molecular genetics of autosomal recessive cerebellar ataxias. *Lancet Neurol.*, **6**, 245–257.
3. Fogel, B.L. (2012) Childhood cerebellar ataxia. *J. Child Neurol.*, **27**, 1138–1145.
4. Shakkottai, V.G. and Fogel, B.L. (2013) Clinical neurogenetics: autosomal dominant spinocerebellar ataxia. *Neurol. Clin.*, **31**, 987–1007.
5. Moreira, M.C. and Koenig, M. (2004, Updated 2011 Dec 8) In Pagon, R.A., Bird, T.D., Dolan, C.R., Stephens, K. and Adam, M.P. (eds), *GeneReviews*™ *[Internet]*. University of Washington, Seattle, Seattle (WA).
6. Anheim, M., Monga, B., Fleury, M., Charles, P., Barbot, C., Salih, M., Delaunoy, J.P., Fritsch, M., Arning, L., Synofzik, M. *et al.* (2009) Ataxia with oculomotor apraxia type 2: clinical, biological and genotype/phenotype correlation study of a cohort of 90 patients. *Brain*, **132**, 2688–2698.
7. Moreira, M.C., Klur, S., Watanabe, M., Nemeth, A.H., Le Ber, I., Moniz, J.C., Tranchant, C., Aubourg, P., Tazir, M., Schols, L. *et al.* (2004) Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat. Genet.*, **36**, 225–227.
8. Ursic, D., Chinchilla, K., Finkel, J.S. and Culbertson, M.R. (2004) Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. *Nucleic Acids Res.*, **32**, 2441–2452.
9. Suraweera, A., Lim, Y., Woods, R., Birrell, G.W., Nasim, T., Becherel, O.J. and Lavin, M.F. (2009) Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. *Hum. Mol. Genet.*, **18**, 3384–3396.
10. Finkel, J.S., Chinchilla, K., Ursic, D. and Culbertson, M.R. (2010) Sen1p performs two genetically separable functions in transcription and processing of U5 small nuclear RNA in Saccharomyces cerevisiae. *Genetics*, **184**, 107–118.
11. Skourti-Stathaki, K., Proudfoot, N.J. and Gromak, N. (2011) Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell.*, **42**, 794–805.
12. Wagschal, A., Rousset, E., Basavarajaiah, P., Contreras, X., Harwig, A., Laurent-Chabalier, S., Nakamura, M., Chen, X., Zhang, K., Meziane, O. *et al.* (2012) Microprocessor, Setx, Xrn2, and Rrp6 co-operate to induce premature termination of transcription by RNAPII. *Cell*, **150**, 1147–1157.
13. Padmanabhan, K., Robles, M.S., Westerling, T. and Weitz, C.J. (2012) Feedback regulation of transcriptional termination by the mammalian circadian clock PERIOD complex. *Science*, **337**, 599–602.
14. Yuce, O. and West, S.C. (2013) Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. *Mol. Cell. Biol.*, **33**, 406–417.
15. Becherel, O.J., Yeo, A.J., Stellati, A., Heng, E.Y., Luff, J., Suraweera, A.M., Woods, R., Fleming, J., Carrie, D., McKinney, K. *et al.* (2013) Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genet.*, **9**, e1003435.

16. Chen, Y.Z., Bennett, C.L., Huynh, H.M., Blair, I.P., Puls, I., Irobi, J., Dierick, I., Abel, A., Kennerson, M.L., Rabin, B.A. *et al.* (2004) DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am. J. Hum. Genet.*, **74**, 1128–1135.
17. Chen, Y.Z., Hashemi, S.H., Anderson, S.K., Huang, Y., Moreira, M.C., Lynch, D.R., Glass, I.A., Chance, P.F. and Bennett, C.L. (2006) Senataxin, the yeast Sen1p orthologue: characterization of a unique protein in which recessive mutations cause ataxia and dominant mutations cause motor neuron disease. *Neurobiol. Dis.*, **23**, 97–108.
18. Fogel, B.L. and Perlman, S. (2006) Novel mutations in the senataxin DNA/RNA helicase domain in ataxia with oculomotor apraxia 2. *Neurology*, **67**, 2083–2084.
19. Criscuolo, C., Chessa, L., Di Giandomenico, S., Mancini, P., Sacca, F., Grieco, G.S., Piane, M., Barbieri, F., De Michele, G., Banfi, S. *et al.* (2006) Ataxia with oculomotor apraxia type 2: a clinical, pathologic, and genetic study. *Neurology*, **66**, 1207–1210.
20. Bernard, V., Minnerop, M., Burk, K., Kreuz, F., Gillessen-Kaesbach, G. and Zuhlke, C. (2009) Exon deletions and intragenic insertions are not rare in ataxia with oculomotor apraxia 2. *BMC Med. Genet.*, **10**, 87.
21. Fogel, B.L., Lee, J.Y. and Perlman, S. (2009) Aberrant splicing of the senataxin gene in a patient with ataxia with oculomotor apraxia type 2. *Cerebellum*, **8**, 448–453.
22. Arning, L., Epplen, J.T., Rahikkala, E., Hendrich, C., Ludolph, A.C. and Sperfeld, A.D. (2012) The SETX missense variation spectrum as evaluated in patients with ALS4-like motor neuron diseases. *Neurogenetics*, **14**, 53–61.
23. Bassuk, A.G., Chen, Y.Z., Batish, S.D., Nagan, N., Opal, P., Chance, P.F. and Bennett, C.L. (2007) In cis autosomal dominant mutation of Senataxin associated with tremor/ataxia syndrome. *Neurogenetics*, **8**, 45–49.
24. Rudnik-Schoneborn, S., Arning, L., Epplen, J.T. and Zerres, K. (2012) SETX gene mutation in a family diagnosed autosomal dominant proximal spinal muscular atrophy. *Neuromuscul. Disord.*, **22**, 258–262.
25. Fogel, B.L. (2011) Interpretation of genetic testing: variants of unknown significance. *Continuum (Minneap Minn)*, **17**, 347–352.
26. Nahas, S.A., Duquette, A., Roddier, K., Gatti, R.A. and Brais, B. (2007) Ataxia-oculomotor apraxia 2 patients show no increased sensitivity to ionizing radiation. *Neuromuscul. Disord.*, **17**, 968–969.
27. Suraweera, A., Becherel, O.J., Chen, P., Rundle, N., Woods, R., Nakamura, J., Gatei, M., Criscuolo, C., Filla, A., Chessa, L. *et al.* (2007) Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. *J. Cell. Biol.*, **177**, 969–979.
28. Duquette, A., Roddier, K., McNabb-Baltar, J., Gosselin, I., St-Denis, A., Dicaire, M.J., Loisel, L., Labuda, D., Marchand, L., Mathieu, J. *et al.* (2005) Mutations in senataxin responsible for Quebec cluster of ataxia with neuropathy. *Ann. Neurol.*, **57**, 408–414.
29. Winden, K.D., Oldham, M.C., Mirnics, K., Ebert, P.J., Swan, C.H., Levitt, P., Rubenstein, J.L., Horvath, S. and Geschwind, D.H. (2009) The organization of the transcriptional network in specific neuronal classes. *Mol. Syst. Biol.*, **5**, 291.
30. Oldham, M.C., Konopka, G., Iwamoto, K., Langfelder, P., Kato, T., Horvath, S. and Geschwind, D.H. (2008) Functional organization of the transcriptome in human brain. *Nat. Neurosci.*, **11**, 1271–1282.
31. Zhang, B. and Horvath, S. (2005) A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.*, **4**, 1–43. Article17.
32. Teglund, S., Olsen, A., Khan, W.N., Frangsmyr, L. and Hammarstrom, S. (1994) The pregnancy-specific glycoprotein (PSG) gene cluster on human chromosome 19: fine structure of the 11 PSG genes and identification of 6 new genes forming a third subgroup within the carcinoembryonic antigen (CEA) family. *Genomics*, **23**, 669–684.
33. Tamaichi, H., Sato, M., Porter, A.C., Shimizu, T., Mizutani, S. and Takagi, M. (2013) Ataxia telangiectasia mutated-dependent regulation of topoisomerase II alpha expression and sensitivity to topoisomerase II inhibitor. *Cancer Sci.*, **104**, 178–184.
34. Bataller, L., Wade, D.F., Graus, F., Rosenfeld, M.R. and Dalmau, J. (2003) The MAZ protein is an autoantigen of Hodgkin's disease and paraneoplastic cerebellar dysfunction. *Ann. Neurol.*, **53**, 123–127.
35. Richard, P., Feng, S. and Manley, J.L. (2013) A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes. Dev.*, **27**, 2227–2232.
36. Coppola, G., Burnett, R., Perlman, S., Versano, R., Gao, F., Plasterer, H., Rai, M., Sacca, F., Filla, A., Lynch, D.R. *et al.* (2011) A gene expression phenotype in lymphocytes from Friedreich ataxia patients. *Ann. Neurol.*, **70**, 790–804.
37. Nishimura, Y., Martin, C.L., Vazquez-Lopez, A., Spence, S.J., Alvarez-Retuerto, A.I., Sigman, M., Steindler, C., Pellegrini, S., Schanen, N.C., Warren, S.T. *et al.* (2007) Genome-wide expression profiling of lymphoblastoid cell lines distinguishes different forms of autism and reveals shared pathways. *Hum. Mol. Genet.*, **16**, 1682–1698.
38. Coppola, G. and Geschwind, D.H. (2012) Genomic medicine enters the neurology clinic. *Neurology*, **79**, 112–114.
39. Fogel, B.L., Lee, J.Y., Lane, J., Wahnich, A., Chan, S., Huang, A., Osborn, G.E., Klein, E., Mamah, C., Perlman, S. *et al.* (2012) Mutations in rare ataxia genes are uncommon causes of sporadic cerebellar ataxia. *Mov. Disord.*, **27**, 442–446.
40. Chen, H., Yang, L., Chen, F., Yan, J., Yang, N., Wang, Y.J., Zhu, Z.R., Hu, Z.A., Sui, J.F. and Hu, B. (2013) Functional inactivation of orexin 1 receptors in the cerebellum disrupts trace eyeblink conditioning and local theta oscillations in guinea pigs. *Behav. Brain. Res.*, **250**, 114–122.
41. Oliver, P.L., Finelli, M.J., Edwards, B., Bitoun, E., Butts, D.L., Becker, E.B., Cheeseman, M.T., Davies, B. and Davies, K.E. (2011) Oxr1 is essential for protection against oxidative stress-induced neurodegeneration. *PLoS Genet.*, **7**, e1002338.
42. Salton, M., Elkon, R., Borodina, T., Davydov, A., Yaspo, M.L., Halperin, E. and Shiloh, Y. (2011) Matrin 3 binds and stabilizes mRNA. *PLoS ONE*, **6**, e23882.
43. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. and Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods*, **7**, 248–249.
44. Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R.M., Blencowe, B.J. and Geschwind, D.H. (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*, **474**, 380–384.
45. Fogel, B.L., Wexler, E., Wahnich, A., Friedrich, T., Vijayendran, C., Gao, F., Parikshak, N., Konopka, G. and Geschwind, D.H. (2012) RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. *Hum. Mol. Genet.*, **21**, 4171–4186.
46. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
47. Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
48. Smyth, G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, 1–25. Article3.
49. Johnson, J.O., Pioro, E.P., Boehringer, A., Chia, R., Feit, H., Renton, A.E., Pliner, H.A., Abramzon, Y., Marangi, G., Winborn, B.J. *et al.* (2014) Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat. Neurosci.*, **17**, 664–666.
50. Piluso, G., Dionisi, M., Del Vecchio Blanco, F., Torella, A., Aurino, S., Savarese, M., Giugliano, T., Bertini, E., Terracciano, A., Vainzof, M. *et al.* (2011) Motor chip: a comparative genomic hybridization microarray for copy-number mutations in 245 neuromuscular disorders. *Clin. Chem.*, **57**, 1584–1596.